# HENP LDRD

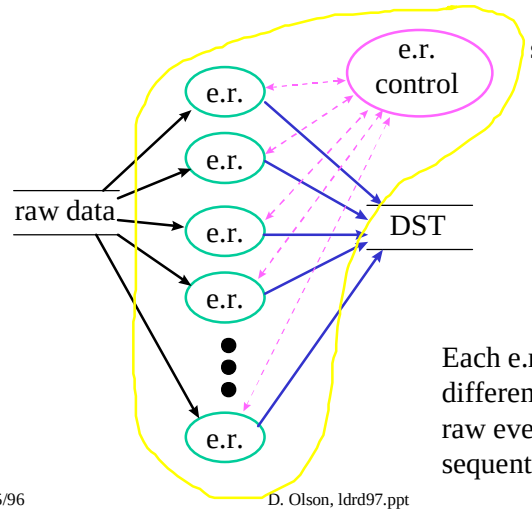
Draft goals, milestones, etc.



# HENP data analysis

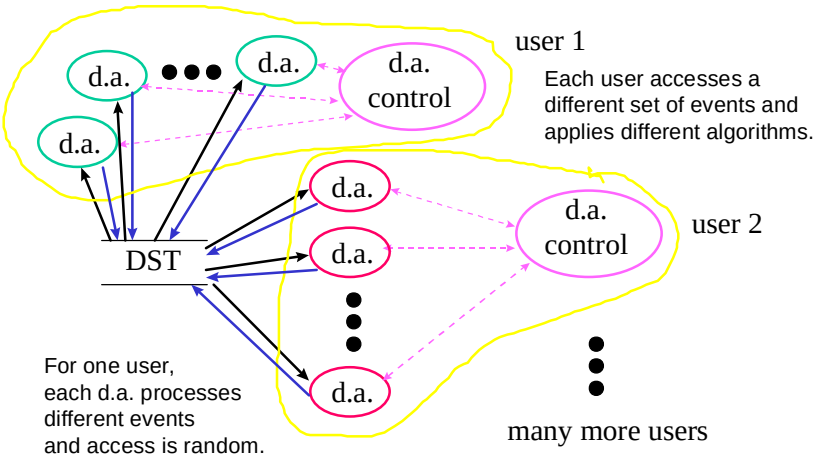
- HENP high-performance data analysis has two primary modes:
  - event reconstruction mode consisting of a single user, sequential data access, parallel event processing
  - analysis mode consisting of many users, random data access, parallel event processing

# Event Reconstruction mode



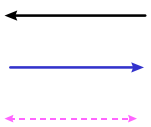



# DST analysis mode



# LDRD implementation

- data read via DPSS
- data write via ?
- control via ? (CORBA?)





# Goals

- Establish a working model for dealing with the data rates and quantities that are used and generated by data analysis and simulations for HENP experiments. The STAR experiment will be the test case of this project. The two analysis modes to be included are an event reconstruction mode (single user, sequential data access, parallel event processing) and a DST analysis mode (many users, random data access, parallel event processing).
- Investigate a model for parallel tape access similar to DPSS disk access.

# Milestones

- 10/96 Brief report describing plan to investigate parallel tape access.
- 10/96 Brief report describing plan for prototype facility that will be used by some real users.
- 11/96 Report on detailed description of requirements for HENP data analysis.
- 12/96 Initial test results of event reconstruction mode.
- 2/97 Initial test results of DST analysis mode with simulated users.
- 5/97 Test results of DST analysis mode with real users.
- 7/97 prototype parallel tape demonstration and characterization.
- 9/97 Final report.



# Deliverables

- 1. A report on the data analysis facility model which includes a roadmap for developing a real production facility that can be supported at NERSC and other sites.
- 2. Quantitative results from testing and monitoring the prototype facility which illustrate the optimum choice of system parameters.
- 3. Prototype facility that is used by real users (physicists) for DST analysis. This usage will provide realistic performance monitoring data.
- 4. Prepare detailed description of requirements for HENP data analysis. This provides input to defining the testing and parameters to monitor in this LDRD project.
- 5. Report on the use of tape as a backing store for random accessed disks.

# Responsibilities

- DPSS interface & testing - Greiman
- Control interface for real users - McParland
- User usage - Olson
- Real simulated STAR data - Olson
- Parallel tape - ?
- Requirements report - Olson (w/ Greiman)
- Final report - Johnston



# Where the $ goes

- Greiman 1.0 FTE
- McParland 0.5 FTE
- Johnston remainder FTE
- hardware (tapes, disks?)

# Questions

- Can we run event processing clients on PDSF?
- What is data write path?
- What is control path?